1976 ✓ (197)

CHANGING CONCEPTS OF CONSCIOUSNESS AND FREE WILL*

R. W. SPERRY†

Introduction

Two special properties of the brain not found in other natural systems, as far as we yet know, have always been notoriously difficult for science to deal with—even in principle. The first of these, of course, is *conscious awareness,* that will-o'-the wisp that science cannot find, cannot demonstrate, measure, or work with and, in most cases, something just the basic nature of which we have been unable to conceive satisfactorily or even imagine. How the brain mechanisms generate subjective conscious experience continues to pose the number one problem for brain research and one of the most truly mystifying unknowns remaining in the whole of science.

The second brain property that science finds particularly troublesome is *free will.* Science is concerned with causal relations and can hardly work out the natural laws, predictions, and understanding of a system that fails to obey the principles of lawful causation. One of the earliest rules for animal behavior stated that, when rigorous conditions are established in which all sensory input can be strictly controlled, one may predict for any measured stimulus that an animal will respond "as it damn pleases." This was widely referred to back in the 1930s as the "Chicago Law of Behavior"—or, in Chicago, as the "Harvard Law."

It is curious and perhaps not entirely coincidental that these same two brain properties that science finds so unaccountable are commonly considered by practically all of us to be the two most important and most treasured of all our brain faculties. When we have lost consciousness, we have lost most of what makes everything worthwhile, and almost the same can be said for the ability to will our own actions, decisions, utter-

*Adapted from a talk presented at York University, Toronto, November 1973, in the Gerstein Lecture Series. The work was aided variously by grant no. MH 03372 from the National Institute of Mental Health, U.S. Public Health Service, the David Stone Foundation, and the F. P. Hixon Fund of the California Institute of Technology.

†Hixon Professor of Psychobiology, Division of Biology, California Institute of Technology, Pasadena, California 91125.

ances and general behavior from moment to moment as we choose. To those of us who look to science for advanced insight and understanding, it always comes as something of a letdown to be reminded that the main strategy science has come up with for dealing with these two most valued and interesting of all the brain properties is, in both cases, simply to ignore or actively to renounce them—to the extent even of expressly denying that either consciousness or free will actually exist as real phenomena. Ever since the advent of behaviorism and adoption of the materialist philosophy in the early 1900s, the prevailing doctrine of twentieth-century science has been telling us that conscious mind and free will are little more than introspective illusions.

According to neuroscience, all brain activity and therefore all behavior is causally determined; and the causal agents and forces are entirely material or physical, that is, biochemical, physiological, electrical, and so on and definitely not mental or anything like the phenomena of subjective experience. A fundamental premise of materialist science holds that a complete explanation of brain function is possible in principle in purely objective physiological and biophysical terms. This objective description and analysis of behavior has seemed to have no need and no place for the likes of inner conscious experience; neither does there seem to be any place apparent in the whole cerebral system where all neural operations are not at all times causally determined. Psychiatry informs us that our slightest slips of tongue, neurotic switches, and other mannerisms and even our dreams can all be shown to have their underlying causes—if one probes deeply enough. In studies with posthypnotic suggestion, it can be shown that actions assumed by the subject to have been carried out spontaneously, of his own free will, were actually in fact preinstructed in detail before witnesses, in a prior hypnotic session that the subject was then ordered to forget.

In other words, in the world view of materialist science, real mental freedom to act and choose is only an illusion, and the whole value-rich world of inner subjective experience gets set aside as some kind of passive, impotent by-product, an epiphenomenal correlate, or just an interior aspect of the one prime material brain process. To be acceptable to science, theories of consciousness have had to be so conceived that it makes no difference to the course of brain events whether consciousness is present or not. Some authorities have preferred to treat the conscious mind as just a spurious pseudo-phenomenon conjured into our thinking by semantic gymnastics, an epistemological artifact.

What remains of the human psyche and the mind of man in the materialist scheme of objective science seems to boil down essentially to a complex system of electrophysicochemical interactions, all causally determined and physically controlled. The resultant view of human nature and the kinds of values that emerge are hardly uplifting. The color,

quality, and higher meaning of life seem to be lost or destroyed; and the long-standing separation between the material world of science, on the one hand, and the world of the humanities and inner experience, on the other, becomes increasingly wide and irreconcilable. The scientific renunciation of conscious mind and free will, flying as it does in the face of common experience and common aspiration, does little to counter the recent waning of intellectual confidence in science and opportunely feeds, instead, various sentiments of antiscience [1].

Attempts to make the behaviorist position more socially acceptable and persuasive, as in the writings of B. F. Skinner [2], prove provocative, perhaps, but hardly compete in appeal with some of the more common humanistic approaches. All of us would prefer to think that we are more than mere puppets of environmental reinforcement and our brain's physiology and that the inner experience we live with most of our waking life is something real and of some material consequence. Many look for further qualities in conscious mind that make possible telepathic communication, psychokinesis, or even precognition, while large numbers would like to count on the properties of conscious awareness to include something immortal and/or reincarnate.

Whatever the kinds of final answers which brain research may eventually lead us to in these areas, there obviously is much at stake. Whole ideologies, world views, philosophies, religious doctrines, and value systems stand or fall on the outcome. If it be true that the materialist world view, as adopted and increasingly supported by science, offers the most valid and accurate interpretation—and certainly science has proven itself again and again against opposing views in the past—then there is much to be said for our facing up to materialist reality and, like behaviorists and the communist world, trying to build our society and life goals and values around and within a philosophy of materialism.

NEW OUTLOOK

That this is indeed the direction in which science must necessarily lead begins to look now not nearly so inevitable as was the case only a dozen years ago. The reason lies in the emergence in recent years of a modified interpretation of the nature of conscious mind and of the fundamental relation of mind to brain mechanism. These latest views represent a substantial swing away from the classic materialist position and give renewed recognition to the role of mental over material forces. A control influence of subjective experience in brain function is predicated within a conceptual explanatory model for psychophysical interaction that is based in terms acceptable to modern neuroscience and does not violate the principles of scientific explanation.

These revised concepts are elaborated in some detail elsewhere [3–5]

and are presented here only in brief outline, in order to put them in perspective and to indicate some of the various implications that emerge. In general, our revised views involve a direct break with established behaviorist doctrine, amounting almost to a full reversal of the central precept on which the behaviorist-materialist movement was founded. Specifically, instead of ignoring or renouncing as nonscientific the subjective conscious mind, the new interpretation recognizes inner conscious awareness to be an integral part of the brain process and a high-level directive force or property in cerebral function. The subjective conscious mind in these revised terms is no longer set aside in the brain of science as a passive correlate or interior aspect but becomes, instead, an essential constituent of higher brain processing. The phenomena of inner experience become active causal determinants in brain activity and are given a functional role and a reason for having been evolved in a physical world.

Since the initial statement of these concepts in 1965 [3], their influence has been apparent with respect particularly to subjectivist approaches in behavioral science. As long as it remained inconceivable that phenomena of conscious experience could affect the course of brain events, those disciplines in psychology that rely on introspective reports of subjective experience, including the clinical, humanist, cognitive, and related schools, continued to be put down in dominant behaviorist thinking as something less than scientific. Once a credible conceptual model for psychophysical interaction became recognized, wherein mental phenomena as top-level controls were neither identical with nor reducible to neural events, the scientific status of consciousness and of the subjective approach underwent a change. Terms like "mental imagery" and visual, verbal, or auditory "images" and all forms of inner thought, motivation, and feeling now became more acceptable as explanatory constructs. After more than 50 years of being strictly avoided on behaviorist principles, such subjective terms have recently exploded into wide usage [6], in a change variously referred to as the "cognitive" [7], "humanist," or "third" [8] revolution, in psychology. Meantime in mind-brain controversy, mentalists, dualists, and psychophysical interactionists have now begun to reappear in force, after having been essentially silent and invisible for decades.

It is not critical at this stage that the new interpretation lacks any firm proof. No proof is available, either, for the behaviorist-materialist position. Just the fact that a scientifically possible explanatory model for psychophysical interaction is conceivable has been sufficient in itself to release the long-pent-up subjectivist pressures.

At the same time, more peripheral movements leaning toward the mystical and supernatural have also been bolstered secondarily in this recent mentalist upsurge, including parapsychology. Actually no direct

support for these latter can be found in our present mind/brain model. If anything, the current interpretation, in which conscious experience becomes a systemic property of and functionally tied to the physical brain process (see below), makes less likely than ever the possible occurrence of mental telepathy, psychokinesis, precognition, and other so-called psi phenomena. Nor can the current view be said to encourage hopes for the existence of any separate, nonphysical realm of conscious mind or spirit divorced from matter. In other words, the current swing away from traditional materialism does not carry us all the way back to dualistic or supernatural concepts but represents, rather, an intermediate compromise within which aspects of both classic materialist and mentalist doctrine are fused in a new combination.

FORMULA FOR PSYCHONEURAL INTERACTION

Without attempting here to review in detail these conceptual developments, let me just restate briefly that, in our current interpretation of consciousness, subjective awareness is conceived to be an emergent property of neural events generated at top levels in the brain hierarchy. The emergent (subjective) properties are conceived to have causal consequences in cerebral activity just as emergent properties commonly do elsewhere. The regulative control role of conscious experience is seen to be based largely in the universal power of any system as a whole over its parts. Mental phenomena built of neural events are conceived to act as dynamic entities in brain organization interacting at their own level in brain function. As high-level dynamic entities, the mental processes control their component biophysical, molecular, atomic, and other subelements, in the same way, for example, that the organism as a whole controls the fate of its separate organs and cells or just as the molecule as an entity carries all its component atoms, electrons, and other subatomic parts through a distinctive time-space course in a chemical reaction. An expanded description of this holistic or entitative type of causal control is presented by Pols [9].

As is the rule for part-whole relations, a mutual interaction between the neural and the mental events is indicated: the brain physiology determines the mental effects, as generally agreed, but the neurophysiology is at the same time reciprocally governed by the higher subjective properties of the enveloping mental operations, as these interact at their own level and prevail upon subsidiary events in brain dynamics. A full causal account of brain function is thus not possible in purely neurophysiological or biophysical terms that do not include these higher, yet-to-be-described mental processes with their subjective pattern properties different from the neural events per se.

By way of illustration, if one could render the nerve impulse and

related glial activity X-ray opaque or radiant and then take fluoroscopic-like pictures of the cerebral turbulence for different kinds of conscious brain events, one should be able in time to begin to describe the critical differences that are responsible, for example, for auditory as opposed to visual or tactual sensation and later to go into further intramodal refinements describing the processing differences involved in seeing red versus green or a triangle versus a square, etc. These conscious processes, as entities, have never been described, and the objective descriptions are still far out of reach. When the objective account becomes available, we will have both the objective and subjective descriptions, but the subjective effect, on these terms, should be understandable and inferrable from the objective description, because the subjective meaning depends on how the brain process, as a dynamic entity, works in the going context of brain activity [10]. The basic organizational features involved are assumed to be genetically determined in very large part.

The foregoing combines important features of both classic dualistic mentalism and monistic materialism. It is mentalistic in that the contents of subjective mental experience are recognized as important aspects of reality in their own right, not to be identified with the neural events as these have heretofore been conceived nor reducible to neural events. Further, the subjective mental properties and phenomena are posited to have a top-level control role as causal determinants [11]. On these terms mind moves matter in the brain. Not only can subjective mind no longer be ignored in science; it becomes a prime control factor in explanatory models. In former theories of consciousness at all acceptable to science, consciousness has been so defined that the causal march of brain mechanisms would proceed the same, whether it is accompanied by subjective experience or not. This is not the case in the present model.

At the same time, the current view can be called materialistic, in that the subjective phenomena as emergent properties of brain activity are built of neural events and therefore always tied, as emergent properties, to the material brain with all its anatomical and physiological constraints. The classic definitions of dualism and monism hardly apply, however, in the current perspective.

FREE WILL VERSUS DETERMINISM

This mind-brain reformulation brings important logical implications also for the interpretation and outlook regarding *free will*. The causal sequence of brain events leading to and determining a given voluntary act or decision no longer is conceived to be restricted to a series of neurophysicochemical activities. The emergent subjective mental properties of these physical processes, as described above, must also be taken

into account and included among the controlling causal determinants. This introduces new degrees and qualities of freedom into the brain's decision-making process, lifting it above the mechanistic, physicalistic kind of determinism envisaged in classical behaviorist, stimulus-response, or materialist doctrine. For example, one's subjective desire to do this or that, along with other subjective feelings and motivations and subjective values of all kinds, plus the whole range of cognitive mental experience, may now, per se, influence the progression of brain events as directive causal factors. As dynamic, holistic properties, the subjective factors are not reducible to, or identifiable with, their neural constituents or as parallelistic correlates of these. In any decision to act, these conscious mental phenomena override and supersede the component physiological and biophysical events involved in the causal progression of brain activity.

A given volitional choice may depend additionally on things like the memory and the mental perspective acquired by the subject (and any consultants) over a span of many decades preceding the decision. Data from the information store of one or more libraries may be called on and funneled into the brain code sequence that leads to the given choice. Even factors like the predicted long-term future consequences of the various alternative choices being contemplated may be included proactively in this vast vortex of cerebral factors that governs the final decision to act. Compared with the kinds of determinism that science deals with in other systems, the degrees and levels of freedom in the operations of the human brain clearly set the brain and mind of man apart with the dignity of an apex post in the universe, far above all other known systems in terms of its ability to choose and to control a course of events.

Even so, one may object that this leaves our brain's decisions nevertheless all, in a sense, determined, even though at this higher, more complex, mental level. We still are caught in the web of a deterministic universe and have to do what we do. Having *degrees* of freedom, in other words, does not quite make for *complete* freedom from causal control. The answer here is that complete freedom from causation would mean behavior based purely on chance, on caprice, and would result in meaningless chaos. What one wants of free will is not to be totally freed from causation but, rather, to have the kind of control that allows one to determine one's own actions according to one's own wishes, one's own judgment, perspective, cognitive aims, emotional desires, and other mental inclinations. This, of course, is exactly what is provided in our current interpretation.

Further Humanistic Implications

These changing concepts of mind and free will substantially alter the general image of man and his role as pictured by science and also bring

other major departures from long-established behaviorist-materialist tradition. When subjective values are conceived to have objective consequences in the brain, they no longer are shut off in a realm outside the domain of science. The old dictum "Science deals with facts, not with values," or "Value judgments lie outside the realm of science," no longer applies in the new framework. That whole dichotomy in which the world view of science is restricted on principle to the cold, objective, quantitative, and materialistic, while all of the subjective, qualitative, and colorful humanistic aspects of life are relegated elsewhere, is negated in the new model.

Old metaphysical dualisms and the seemingly irreconcilable paradoxes that formerly prevailed between the realities of experimental brain science, on the one hand, and those of the inner experience, on the other [12], become reconciled today in a single, comprehensive, and unifying view of mind, brain, and man in nature. We pass conceptually in a single continuum within the brain from the brain's lower subnuclear particles on up through atoms, molecules, cells, and successively higher levels to brain circuit systems without consciousness, on into those special, high-order cerebral processes with subjective properties.

Instead of separating science from values, the present interpretation (when all the various ramifications and logical implications are followed through) leads to an almost diametrically opposed philosophy, supporting science as man's prime hope in the search for a new ethic and higher meaning. Science emerges as the most effective and valid means by which the human brain can approach comprehension of the forces that move the universe and created man. From the total collective knowledge of all the sciences, we get the best world-view framework for determining ultimate meaning and for finding ethical axioms and guideline beliefs to live and govern by.

The Conscious Self Bisected

Rather than elaborate further on all this humanistic fallout—and I am afraid that this is just the beginning of a whole long story [13], I plan to turn back now to some of the research evidence itself and to look at some of the observations and kinds of data that have helped in recent years to steer our thinking in these directions. First, let me review quickly some of the effects of dividing the brain surgically down the midline, including results that seem to show that conscious experience is tied not only to the living active brain but, more specifically, to certain neural systems at the upper cerebral levels and that the normal unity of the inner world of consciousness is subject to midline division by surgery. When the fiber cables that link right and left cerebral cortex, including the enormous corpus callosum with its 200,000,000 or so lines of cross communication,

are surgically sectioned in animals or in human patients with advancing intractable epilepsy, the effects at the level of conscious activity are striking. Careful testing of the independent function of each hemisphere separately suggests that the unified world of inner experience is also divided into two separate right and left systems, each hemisphere apparently conscious within itself but unaware of the perceptual, learning, and related memory experiences of its partner within the same cranium. The normally unified perception of the field of vision, for example, is divided into two separate, inner, visual realms, one in each of the disconnected hemispheres and each cut off from direct intercommunication.

Further, the surgically separated hemispheres in man have been found to utilize different modes of perceptual and cognitive processing. Subtle left-right differences in cognitive style, not readily apparent otherwise, can be demonstrated to advantage after cerebral commissurotomy. Direct comparisons can be made for performance on the same task, in the same individual, where the life history and other background factors are all equated and where the two hemispheres can be pitted against each other on equal terms for response dominance. Fine differences can be measured under these conditions and qualitative distinctions discerned that are much more difficult or even impossible in clinical patient comparisons involving different persons. One hardly takes notice if different people with brain damage tend to think a little differently. Even subtle differences become meaningful, however, when you see the same individual performing the same test tasks and reaching the same solutions in two different ways, using different strategies, much like two different people, depending on whether it is the left or the right hemisphere in use [14, 15].

Repeated examination has consistently confirmed the strong lateralization and dominance for speech, writing, and calculation in the disconnected left hemisphere in right-handed patients. The minor, right hemisphere, by contrast, is unable as a rule to respond in speech or writing, although it comprehends slowly spoken instructions quite well. Nor can it typically perform calculations, except for simple additions up to sums less than 20. The language-dominant hemisphere is the more aggressive, executive, leading hemisphere in the control of most behavior. After surgery, these patients seem to run primarily on the left hemisphere that we mainly see in action and the one with which we regularly communicate. It is the highly developed linguistic, logistic, and related cognitive capacities of the left hemisphere, apparently, that are largely responsible for earlier impressions that cerebral functions persevere with little impairment in the absence of the commissures.

The disconnected mute, "minor" hemisphere, by contrast, seems to be carried along in most situations as a largely passive, silent passenger

obliged to leave the behavioral driving mainly to its partner. As a result, the nature and quality of the inner mental world of the nonvocal right hemisphere remains relatively inaccessible to investigation, requiring special testing measures with nonverbal forms of expression. Although some authorities have been reluctant to credit the disconnected minor hemisphere even with being conscious, it is our own interpretation, based on a large number and variety of nonverbal tests, that the nonvocal hemisphere is indeed a conscious system in its own right, perceiving, thinking, remembering, reasoning, evaluating, willing, and emoting, all at a characteristically human level, and that both the left and the right hemisphere may be coconscious simultaneously in different, even in mutually conflicting mental experiences that run in parallel.

Although predominantly mute and generally inferior in all performances involving language or linguistic, mathematical, or sequential reasoning, the right hemisphere is nevertheless clearly the superior cerebral member for certain types of tasks. These are, of course, all nonlinguistic, nonmathematical functions. Largely, they involve the apprehension and processing of spatial patterns, relations, and transformations. They seem to be holistic and unitary rather than analytic and fragmentary, orientational more than focal, and seem to involve concrete perceptual synthetic insight rather than symbolic or feature-by-feature sequential reasoning like the left [16].

The two disconnected hemispheres exhibit such independent properties and distinct mental capacities that we come to think of each hemisphere as having a separate mind of its own. Resultant inferences about the nature and locus of consciousness have ranged rather widely [17]. At one extreme it has been inferred that, even in the normal intact brain, only the left hemisphere contributes conscious experience. At the opposite extreme, others infer that, in the normal intact state, each of the hemispheres has its own separate domain of conscious experience and that each of us therefore is really a compound of two separate left and right persons, which becomes evident only after commissurotomy. Between these extremes, several intermediate possibilities have also been recognized [18].

Our own view of the evidence has favored the conclusion that both hemispheres in the surgically disconnected state are separately conscious, with a few qualifications [15], and that, in the normal intact state, the conscious activity is typically a unified and coherent bilateral process that spans both hemispheres through the commissures. This latter implies that the connecting fiber systems of the brain help to mediate subjective experience as well as the switching sites and transmission interfaces of the cerebral neuropile. It also assumes that the commissural fiber systems interconnecting the hemispheres are not different in principle, in this respect, from those making connections within a hemi-

sphere. This interpretation is based also on a theory of consciousness that goes back to the early 1950s [10], in which the subjective sense of unity in conscious experience and subjective meaning generally are conceived to derive not so much from the particulate infrastructure of the given neural process or from its continuity or its form as an inner copy of the perceived object but, rather, from its operational effect or the way it works in the context of brain dynamics. This latter in turn will be seen to imply further a causal influence which brings us full circle to points raised at the outset.

REFERENCES

1. T. Roszak. Where the wasteland ends. Garden City, N.Y.: Doubleday, 1973.
2. B. F. Skinner. Beyond freedom and dignity. New York: Knopf, 1971.
3. R. W. Sperry. Bull. At. Sci., **22**:7, 1966.
4. R. W. Sperry. Psychol. Rev., **76**:532, 1969.
5. R. W. Sperry. Psychol. Rev., **77**:585, 1970.
6. Z. W. Pylyshyn. Psychol. Bull., **80**:1, 1973.
7. W. N. Dember. Am. Psychol., **29**:161, 1974.
8. F. W. Matson. *In:* P. Zimbardo and C. Malach (eds.). Psychology for our times, p. 19. Glenview, Ill.: Scott, Foresman, 1973.
9. E. Pols. Int. Philos. Q., **11**:293, 1971.
10. R. W. Sperry. Am. Sci., **40**:291, 1952.
11. R. W. Sperry. *In:* G. G. Globus, G. Maxwell, and I. Savodnik (eds.). Consciousness and the brain, chap. 6. New York: Plenum, 1976.
12. T. W. Wann (ed.). Behaviorism and phenomenology. Chicago: Univ. Chicago Press, 1965.
13. R. W. Sperry. Perspect. Biol. Med., **16**:115, 1972.
14. J. Levy. *In:* S. Dimond and J. Beaumont (eds.). Hemisphere function in the human brain, p. 121. London: Elek, 1975.
15. R. W. Sperry. *In:* F. O. Schmitt and F. G. Worden (eds.), The neurosciences third study program, p. 5. Cambridge, Mass.: M.I.T. Press, 1974.
16. J. Levy, C. Trevarthen, and R. Sperry. Brain, **95**:61, 1972.
17. O. L. Zangwill. *In:* S. Dimond and J. Beaumont (eds.) [14], p. 264.
18. T. Nagel. Synthese, **22**:396, 1971.